

Effect of observational holes and borders on lacunarity and fractality measurements in a galaxy catalogue

J. E. García-Farieta,^{1*} R. A. Casas-Miranda,¹

¹*Departamento de Física, Universidad Nacional de Colombia - Sede Bogotá, Av. Cra 30 No 45-03, Bogotá, Colombia*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Cosmological observations reveal that the Universe contains a hierarchy of galaxy clustering with a transition to homogeneity on large scales according to the Λ CDM model. Some observational estimates suggest that the Universe behaves as a multifractal object, where galactic clustering is based on generalisation of the dimension in metric spaces. From this point of view, we study the spatial distribution of points by simulating galaxies on large scales in the Universe with samples from the Sloan Digital Sky Survey (SDSS), including observational holes in the masks. We build homogeneous samples following the radial selection function using the “shuffle” method for a main sample of 3,273,548 points limited to the redshift range $0.002 < z < 0.2$. A random distribution of observational holes in right ascension and declination was created from the SDSS-Baryon Oscillation Spectroscopic Survey (BOSS) footprint. We determined the fractal dimension $D_q(r)$ in the range $-6 < q < 6$ and the lacunarity spectrum using the sliding window technique to characterise the hierarchical clustering in these catalogues. Our results show that observational holes cause a shift in the homogeneity scale; for $q > 0$ and percentages of holes near 40%, r_H is displaced on scales on the order of $120 \text{ Mpc}/h$. Hole percentages between 10% and 30% show an r_H of $70 - 90 \text{ Mpc}/h$, and for percentages below 10%, r_H decreases to become equal to the r_H value of the SDSS-BOSS footprint `boss_survey.ply`. For $q < 0$, the homogeneity scales have strong fluctuations for all hole percentages studied.

Key words: large-scale structure, galaxy clustering, observational holes, multifractal analysis

1 INTRODUCTION

The large-scale structure of the Universe has been extensively studied since the publication of the first galaxy surveys. Some analysis with high-redshift samples shows the existence of fractal correlations (Celerier & Thieberger 2005). The large network of matter filaments that connect galaxies is characterised by fractal structures (Martinez & Saar 2010). In a cosmological context, fractals were introduced into a model by Mandelbrot as a hypothesis to solve the Olbers paradox (Martínez & Jones 1990). According to the fractal hypothesis, if the set of stars forms a structure with self-similarity properties, such as a Cantor dust, the paradox is solved, for even in a mathematically infinite universe, the sky can have dark regions. We currently know that at small scales, the galaxy distribution exhibits fractal behaviour (Peebles 1980; Mandelbrot 1982). Some observational estimates suggest that the Universe behaves as a mul-

tifractal object with a Hausdorff dimension $D_H = 2.1 \pm 0.1$ and correlation dimension $D_2 = 1.3 \pm 0.1$ (Durrer & Labini 1998).

The introduction of the fractal concept allows interpretation of the hierarchical clustering in terms of self-similarity properties or an invariant scale of the galaxy distribution (Chacón 2014). It seems clear that for $r < 10^{-1} \text{ Mpc}$, galaxy clustering exhibits fractal behaviour (Peebles 1993); at larger scales, the structure is more complex and can be characterised as a multifractal object, requiring a generalisation of the fractal dimension in order to explain this distribution on the same scale range (Conde-Saavedra et al. 2015).

Fractal analysis is a useful mathematical tool that quantifies galactic clustering using data from galaxy surveys by calculating quantities such as the fractal dimension, making it possible to establish relationships between these values and other statistical descriptors. The possible cosmological implications of fractal analysis of the galaxy distribution are discussed in detail in (Baryshev et al. 1998; Martinez 1991).

* E-mail: joegarciafa@unal.edu.co

Because cosmic clustering developed under the influence of gravity alone, there is no physical motivation to consider smaller scales related to the processes of galaxy formation. In this context, the multifractal measures are described as a natural scaling result of the matter interactions in the Universe. A formal and general proof of the fractal nature of the matter distribution on large scales requires additional observations and some solutions of the Einstein field equations that take into account the conditional cosmological principle (Martínez & Jones 1990). An additional motivation to study the matter distribution on large scales from a multifractal viewpoint is the transition to a homogeneity scale, which can be defined according to (Yadav et al. 2010) as the value of r above which the fractal dimension D_q of the distribution is equal to the dimension of the physical space in which the points are distributed, i.e., $D = 3$; this implies that if the distribution is finite in size and weakly clustered, such as that in the first galaxy catalogues, it is difficult to achieve this equality. For a non-integer dimension, the galaxy distribution is in a state of transition to homogeneity (Grujić & Panković 2009).

The clustering description under the fractal hypothesis has encouraged the development of several theoretical models. One of the first models was proposed by Mandelbrot (1982) using Lévy flights to simulate the distribution of galaxy clusters regardless of the underlying physical phenomena. Similarly, Peebles (1989) developed the basis for a statistical description of structures in the Universe; using correlation functions of up to four points for the first galaxy catalogues, he showed that the fractal dimension depends on the size of the distribution. Hence, this quantity changes from a pure fractal with dimension $D = 1.26$ for $r < 15.33 \text{ Mpc}$ to structures consistent with the standard cosmological principle on larger scales (Peebles 1989). Some observations have shown that the galaxy distribution grows in proportion to r raised to an exponent related to the correlation dimension. For the galaxy distribution on a significant range of scales, this dimension is approximately 2 (Martínez & Saar 2010), a very different value from 3, which is what would be observed for a homogeneous distribution because the number of galaxies grows in proportion to the volume of a sphere, as expected according to the standard cosmological model given the homogeneity and isotropy conditions.

Recent studies have also reported fractal and multifractal behaviour of the spatial distribution of galaxies and dark matter halos using galaxy surveys, in which large-scale fractal structures are evident (Pino et al. 1995); in particular, see (Zheng et al. 1988) for *CfA survey data*, (Xia et al. 1992) for the *IRAS survey*, and (Seshadri 1999) for the *Campanas redshift survey*. Chacón-Cardona & Casas-Miranda (2012) performed a multifractal analysis of dark matter halos from the *Millennium* simulation (Springel et al. 2005); they found a transition to homogeneity between 100 and 120 Mpc/h , in strong agreement with the ΛCDM model. In contrast, an analysis with volume-limited samples from SDSS-DR7 (Chacón 2014; Chacón-Cardona et al. 2016; Muñoz-Cuarteras & Mueller 2012) reported fractal behaviour at large scales until 165 Mpc/h , where D_q is always below the homogeneity limit $D = 3$ for all values of the structure parameter. This result is consistent with (Joyce et al. 1999; Labini et al. 2009), who found hierarchical patterns with self-similarity properties and a fractal dimension smaller than the dimen-

sion of physical space at scales greater than 100 Mpc/h . For fractal analysis with WiggleZ, (Scrimgeour et al. 2012) reports a transition to homogeneity at $r_H = 71 \pm 8 \text{ Mpc}/h$ with $z \leq 0.2$; this indicates that the galaxy distribution does not behave as a fractal object. This result is also consistent with those of (Hogg et al. 2005; Yadav et al. 2005; Sarkar et al. 2009), who report a transition to homogeneity at $\sim 70 \text{ Mpc}/h$. Wu et al. (1999) and Yadav et al. (2005) are in agreement with this value for the homogeneity scale; however, their studies show that at smaller scales the galactic cluster has fractal properties with dimension $D \approx 1.2 - 2.2$. In all cases, the effects of the geometry of the surveys must be taken into account according to (Yadav et al. 2010); in particular, the fractal calculations may be affected by the presence of holes and borders in the catalogues that are inherent to the process of observation using astronomical instruments. In this paper, we studied the multifractal behaviour of a spatial distribution of points using homogeneous samples built from the Sloan Digital Sky Survey-Baryon Oscillation Spectroscopic Survey (SDSS-BOSS) footprint `boss_survey.ply`, including a random distribution of observational holes in right ascension and declination. We determined the fractal dimension in the range $-6 < q < 6$, the lacunarity spectrum, and the homogeneity scale using the sliding window technique for each sample.

In section 2, we define the main concepts of fractal formalism, such as the fractal dimension, multifractal description of galaxy clustering, lacunarity, and generalised fractal dimension. In section 3, we review the origin of observational holes, borders, and footprints of SDSS masks. Moreover, we present the construction of synthetic homogeneous samples limited in redshift, including the distribution of observational holes. Then, in section 4, we present the results of the application of our method to the constructed samples with the BOSS footprint. In section 5, the results for the multifractal dimension and lacunarity are discussed. Finally, our conclusions are given in section 6.

2 FRACTAL AND MULTIFRACTAL FORMALISM

2.1 Concepts of fractal dimension

The concept of dimension can be associated with the number of degrees of freedom or the minimum number of coordinates to specify any point within a distribution of points in a metric space. Topologically, the dimension indicates how much space a set occupies near each of its points (Falconer 2004). The most intuitive definition of dimension is the *topological dimension* D_T , introduced by Poincaré and generalised by Lebesgue. In this definition, given a set of topological space $X \in \mathbb{R}^n$, the dimension is the minimum value of n for which every open cover admits a locally finite open refinement the order which does not exceed $n + 1$. If there is not a minimum value of n , we say that the set is infinite-dimensional. In Euclidean space, \mathbb{R}^n , D_T takes integer values in $(0, n)$ (Mandelbrot 1982).

When the sets describe irregular shapes, the concept of dimension in terms of the number of coordinates is insufficient to describe them. This fact has motivated the introduction of new concepts beyond the classical geometry (Mandelbrot

1982); therefore, fractal geometry was developed. Thus, it is possible to give a different concept of *dimension*; for instance, the *self-similar dimension* can be explained by further fragmentation of an object or set, and the ratio of the number of identical parts, where each part is scaled down by the ratio r . For any set $X \in \mathbb{R}^n$ that supports division into a finite number of subsets $N(k)$, where all of them are consistent with each other by translations and rotations, and it is a reduced copy of the initial set by a factor $r = 1/k$, the self-similar dimension of X is defined as the unique value D satisfying the equation $N(k) = k^D$ (Mandelbrot 1982), i.e.,

$$D = \frac{\log N}{\log(k)}. \quad (1)$$

Here D is not necessarily an integer number; in some cases, it may be an integer and match the topological dimension. When the object cannot be subdivided into exact copies of itself, we can use the *box-counting dimension*, in this case, a set \mathcal{A} covered by a grid or regular boxes, all equal to each other, with side $\delta > 0$; then, the number of boxes $N(\delta)$ needed to cover the figure is determined. This process is very natural for a computer, and it is not necessary that the figure be self-similar. Hausdorff and Besicovitch (Mandelbrot 1982) proposed a more general definition of *dimension* that considers fractional values and can be defined for any set of points. Mandelbrot defined a fractal object as a set with Hausdorff dimension D_H strictly exceeding its topological dimension; thus, sets with non-integer Hausdorff dimension are fractals. In practice, the Hausdorff dimension is not always easy to calculate (Falconer 2004). In this case, following the idea of the *self-similar fractal dimension*, the *mass-radius fractal dimension* D_m is defined by a power law; this dimension is the measure of the total mass contained in a sphere of radius R whose center is a point of the set, and the mass contained as a function of the radial size is determined as

$$M(R) = FR^{D_m}, \quad (2)$$

where the factor F is a function that may be different for fractals with identical dimension, and the density number of galaxies decreases as $n(R) \approx R^{D_m-d}$ for a set in \mathbb{R}^d (Martínez & Saar 2010).

2.2 Multifractals, lacunarity, and generalised fractal dimension

For point distributions from galaxy catalogues, the analysis is more complex than that for modelled distributions because fluctuations associated with the intrinsic characteristics of the mass distribution appear, so we need to use more general geometric estimators (Saslaw 2000). The fractal dimension is not sufficient to characterise a set unambiguously; consequently, the concept of multifractals and lacunarity have been introduced (Seshadri 1999).

Multifractals provide the most detailed description possible of the fractal properties of a distribution of points. Given a distribution in which each region exhibits a fractal behaviour, but the dimension changes from one place to another, it is possible to establish correlations between these dimensions and to do a complete analysis from a higher level (Blumenfeld & Mandelbrot 1997). The lacunarity is an estimator associated with the size of the holes and borders

within the distribution; furthermore, it describes the “texture” of a fractal. This concept can be extended by other definitions of the fractal dimension as a correlation dimension (Martínez & Saar 2010). Generally, if a fractal has large empty regions within the clustering, i.e., “lacunae” or holes, it will have a high lacunarity value, which increases the heterogeneity of the structure, and if a fractal is translation-invariant, the lacunarity will be low, and the structure will be very close to homogeneous. Different fractals can have the same dimension but very different lacunarity values.

For physical properties that depend on the scale, the fractal behaviour at small scales reported by Peebles (1989), Bagla et al. (2008), and Martínez & Jones (1990) can be extended to a matter distribution at large scales using the multifractal formalism. To determine the lacunarity, first we must know the F factor in equation (2), which is related to the average distance between neighbours. For each center of the point distribution, the number of particles $n_i(r)$ contained within a sphere of radius r measured from the position of the i th particle is given by

$$n_i(r) = \sum_{j=1}^N \Theta(r - |\mathbf{x}_i - \mathbf{x}_j|), \quad (3)$$

where the sum is over all particles in the sample, and N is the total number of particles. The coordinates of each particle in the three-dimensional space are denoted as \mathbf{x}_j , $j \neq i$, and Θ is the Heaviside function. The number of particles $n_i(r)$ around each galaxy taken as the center, with coordinates \mathbf{x}_i , is determined by counting the particles around the center that are located within a comoving sphere of radius r (Celerier & Thieberger 2005).

The correlation dimension is defined similarly to the mass-radius fractal dimension (Seshadri 2005). To characterise the distribution, we must have all the information about the statistical moments in order to define the generalised dimension. The generalised correlation integral $C_q(r)$ is defined as

$$C_q(r) = \frac{1}{NM} \sum_{i=1}^M [n_i(r)]^{q-1}, \quad (4)$$

where q is called the structure parameter and corresponds to an arbitrary real number, and M is the number of particles used as centers. According to (Murante et al. 1997), from the correlation integral it is possible to perform a power series expansion of $\log(r)$ [equation (5)] and thus to calculate directly the multifractal dimension D_q and the lacunarity spectrum Φ_q . It is sufficient to keep the first two terms on the right side of equation (5), which is simplified so that a simple relation between the generalised correlation integral and generalised fractal dimension is obtained as in equation (6) (Chacón 2014).

$$\log[C_q^{1/(q-1)}] = D_q \log(r) + \log(F_q) + \mathcal{O}\left(\frac{1}{\log(r)}\right). \quad (5)$$

$$C_q(r)^{1/(q-1)} = F_q r^{D_q}. \quad (6)$$

Thus, the generalised fractal dimension and generalised lacunarity can be defined in the same way as the mass-radius fractal dimension. The generalised dimension D_q is given by

$$D_q = \frac{1}{q-1} \frac{d \log C_q(r)}{d \log r}, \quad (7)$$

and the generalised lacunarity, using the previously determined factor F for each structure parameter q , will be

$$\Phi_q = \frac{\langle (F_q - \langle F_q \rangle)^2 \rangle}{\langle F_q \rangle^2} = \frac{\langle F_q^2 \rangle}{\langle F_q \rangle^2} - 1. \quad (8)$$

For some values of q such that $q_i \neq q_j$ which satisfy $D_{q_i} = D_{q_j}$, i.e., D_q is independent of q and r , the distribution is called monofractal because its dimension is constant. In addition, if D_q is equal to the Euclidean dimension, the distribution is homogeneous. For $q \geq 1$, D_q explores the scaling behaviour in high-density environments within the distribution, which are associated mainly with clusters and superclusters, whereas for $q < 1$, D_q explores the behaviour in low-density environments, i.e., those associated with *voids* (Sarkar et al. 2009). A full spectrum of the generalised fractal dimension provides detailed information about the entire distribution, whether in regions of high density or low density. This allows us to connect the concept of fractal dimension with statistical measures used to quantify the distribution of matter on large scales. If the distribution of galaxies undergoes a transition to homogeneity, all values of the fractal dimension tend to the dimension of the physical space, that is, $D_q \simeq D = 3$ for any value of q ; at small scales, we expect to see a spectrum of values of the fractal dimension all different from 3, as strongly structures defined before of a homogeneity transition; in this case, the lacunarity spectrum must approach zero at the same radial scale r .

3 DATA SAMPLES

3.1 Observational holes, borders, and footprint

Galaxy surveys that use an optical fiber multi-object spectrograph require automated methods to controlling each fiber and effectively determining the regions of interest to record information from the target. In SDSS-I/SDSS-II, the fibers capture light from 640 objects simultaneously; in SDSS-III, this value increased to 1000. In both cases, the objects were observed in a circular field of radius 1.49 called “tile” (Blanton et al. 2003). Although automated mechanisms were used, there are effects that cause erroneous or unmapped regions, which result in holes in the catalogue. Specifically, these effects are associated with the position of the optical fibers; the most significant of these effects is fiber collisions, which occur when two objectives are close enough together in the same observation that the fibers cross each other, exceeding the allowed angle. In SDSS-I/II, the collision radius was $55''$; in SDSS-III BOSS, it was $62''$. There are additional effects inherent to the process of observation that cannot be ignored and that also produce sky regions that should be eliminated from the samples for various reasons, including effects due to seeing, bright stars that saturate the detectors, trails caused by meteors and satellites or nearby objects, and failures in the observation plates. These effects are identified and quantified in detail using masks.

The accumulation of these effects results in a distribution of observational holes in the galaxy samples. The effective area covered by the catalogue depends on the masks’ quality and the sectors to be excluded from observation. The various types of masks depend exclusively on the catalogue. SDSS-I/II has the following masks: *Bleeding*, *Bright_star*, *Trail*,

Hole, and *Seeing*. SDSS-III uses *Bright Star Mask*, *Center-post Mask*, *Bad Field Mask*, and *Collision Priority Mask* [for details on each mask, see (Blanton et al. 2001; Smeed et al. 2012; Dawson et al. 2013)]. All of these masks are exclusion masks and indicate the regions in which the data quality is unacceptable; i.e., if a random point is within the mask, it will be excluded from the catalogue. According to (Blanton et al. 2001), these areas are relevant to and useful for large-scale structure studies. For fractal analysis in particular, Yadav et al. (2010) refers to some effects that may be caused by the geometry of the catalogues, as holes not only produce incompleteness regions, but also modify the geometry of the masks by adding edges or borders.

Although masks include observational holes, there is a footprint, i.e., an observational template with the SDSS geometry, that represents the ideal galaxy mapping without holes, and it covers the largest area possible. In this paper, we used the BOSS footprint `boss_survey.ply`. The masks and SDSS footprint are designed in convex polygons that are independently generated for each of the five filters in the catalogue. The geometry of the masks is given in terms of spherical polygons and arrays; these files are manipulated using the *Mangle* code (Hamilton & Tegmark 2004). In this case, a mask is an arbitrary union of weighted angular regions bounded by a number of edges.

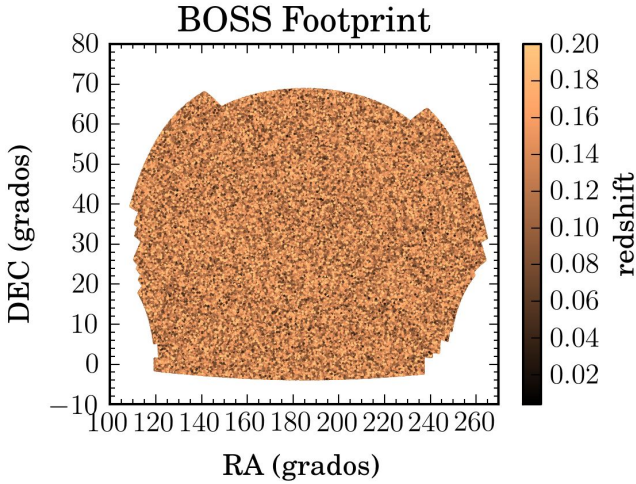
3.2 Synthetic samples

The synthetic samples have a random point distribution that represents galaxies observed by the SDSS; they were built on the basis of the BOSS *footprint* to cover the largest possible observation area, corresponding to 14555 deg^2 . The BOSS footprint is made up of 19 polygons for the two hemispheres of the catalogue. In this paper, we used only the northern hemisphere from the spectroscopic catalogue to cover $\sim 9399 \text{ deg}^2$ (see Fig. 1). The southern hemisphere polygons were removed from `boss_survey.ply`. Then absolute equatorial coordinates $(\alpha, \delta) \equiv (RA, DEC)$ were assigned to each point within the *footprint*. First, we selected a random polygon from the mask with a probability proportional to the product of the statistical weight and area of the polygon; this was then limited, and a random point was generated within the circle. We checked whether the point was located within the polygon. If it was, the point was kept; otherwise, the algorithm was repeated. The number density of points generated in the footprint corresponds to 3,273,548, the same value as in the random catalogue `random0_DR11v1_LOWZ_North.dat` available in SAS <http://data.sdss3.org/sas/> (see Table 1). The *random* catalogues have the same structure observed in SDSS-DR11 `galaxy_DR11v1_LOWZ_North.dat`, with 2,401,952 galaxies in the redshift range $0.00014 < z < 0.98723$.

Synthetic catalogues based on the BOSS footprint were limited in redshift to $0.002 < z < 0.2$, so the time interval between two events at these points is sufficiently small, and there are not significant changes in the large-scale structure patterns (Laporte 2014). To model the redshift distribution of galaxies, it is necessary to have a complete theoretical model that takes into account the behaviour of the population of galaxies. Although a complete model is not available (Ross et al. 2012), the effects of the galaxy distribution in redshift can be known from information provided

Table 1. Number density in the SDSS galaxy survey, random and synthetic samples.

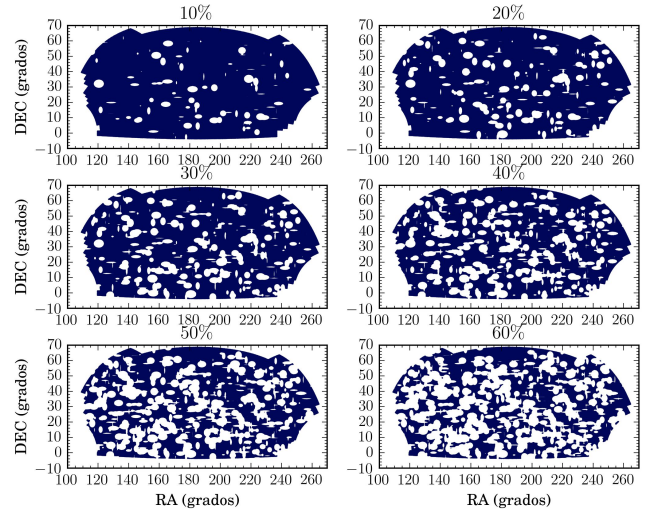
Density Number	DR7		DR10-LOWz		DR11-LOWz		BOSS Footprint Shuffler
	Random	Galaxy	Random	Galaxy	Random	Galaxy	
Unlimited sample $0.00014 < z < 0.98723$	1925442	105832	10162336	203614	14035192	281103	3273548
Limited sample $0.002 < z < 0.2$	1765923	97064	2329271	49420	3273548	66314	


Figure 1. North Galactic projection in absolute equatorial coordinates from BOSS footprint `boss_survey.ply` spectroscopic catalogue generated using *Mangle*. Colour bar indicates the redshift distribution limited to $0.002 < z < 0.2$.

by the observed catalogues, some cosmological simulations, and random catalogues, considering that they always follow a radial selection function $n(z)$ (Ross et al. 2012; Kazin et al. 2010). Naturally, the samples will be affected by the effects of the radial selection function, which result from the instrumental inability to detect faint galaxies at great distances. To determine $n(z)$, we used the shuffle method suggested by Ross et al. (2012) and Kazin et al. (2010), which consists of assigning randomly to each point of the sample a redshift value selected from the galaxy catalogue.

To build homogeneous synthetic samples, first we filtered the redshift values of the DR11 galaxy catalogue on $0.002 < z < 0.2$. Next, we used the shuffle method to assign the radial selection function and distribute all the points in the BOSS footprint to complete the synthetic sample.

For each point in the synthetic samples, the coordinates $(RA, DEC, z) \equiv (\alpha, \delta, z)$ are known. The area of observational holes in the SDSS-DR10 and DR11 masks has a maximum size of approximately 36.727 deg^2 ; these holes can be modelled as ellipses with a semi-major axis $a_{max} = 3.893^\circ$ and semi-minor axis $b_{max} = 3.003^\circ$. The footprint holes are generated randomly by assigning a random point with coordinates $(\alpha_{rnd}, \delta_{rnd})$; an ellipse centered at this point is drawn with semi-axes $a \in (0, 3.893]^\circ$ and $b \in (0, 3.003]^\circ$, and then the ellipse is removed from the footprint. This process is done for different percentages of holes based on the total number of points in the mask. Thus, we obtain the samples *footprint-10*, *footprint-20*, *footprint-30*, *footprint-40*,


Figure 2. $\alpha - \delta$ projections of the samples with holes limited to $0.002 < z < 0.2$. The percentages of holes were calculated on the basis of the total number of data points from the DR11 random sample.

footprint-50, *footprint-60*, and *footprint-70* for the corresponding percentages. All the samples with holes are correlated because they are built cumulatively, i.e., *footprint-10* contains *footprint-20*, which contains *footprint-30*, and so on. In Fig. 2, the projections in $\alpha - \delta$ are illustrated for the samples used in the multifractal analysis described in the next section.

4 MULTIFRACTAL ANALYSIS AND LACUNARITY SPECTRUM

In the previous section, we described how we built the limited synthetic samples with $0.002 < z < 0.2$. The homogeneous footprint without holes is the densest sample in terms of the number of points, with the positions of 3,273,548 objects in absolute equatorial coordinates (see Table 2). The observational hole distribution in the samples covers values from 10% to 70% in the northern hemisphere BOSS footprint and contains an ideal sampling of all points registered in the catalogue with a total area of $\sim 9399 \text{ deg}^2$ (see Fig. 2). The multifractal analysis was done in comoving Cartesian coordinates in order to compare the results with the standard cosmology.

Moreover, to perform the multifractal analysis, we made a sampling within the limited redshift samples without any assumptions about the shape of the group of points; consequently, it is necessary to determine the radius of a larger

Table 2. Summary of redshift-limited samples created from the BOSS footprint with different percentages of observational holes used in the multifractal analysis.

Sample	Area Ω (deg ²)	Area Ω_{holes} (deg ²)	R_s (Mpc/h)
0%	9399	0	396.68
10%	8459	940	
20%	7519	1880	
30%	6579	2820	
40%	5639	3759	
50%	4699	4699	
60%	3759	5639	
70%	2820	6579	

sphere than these can contain. According to [Gabrielli et al. \(2006\)](#), the effective depth R_s for all samples according to their limits in $\alpha - \delta$ can be determined by equation (9).

$$R_s = \frac{R_d \sin(\delta\theta/2)}{1 + \sin(\delta\theta/2)}, \quad (9)$$

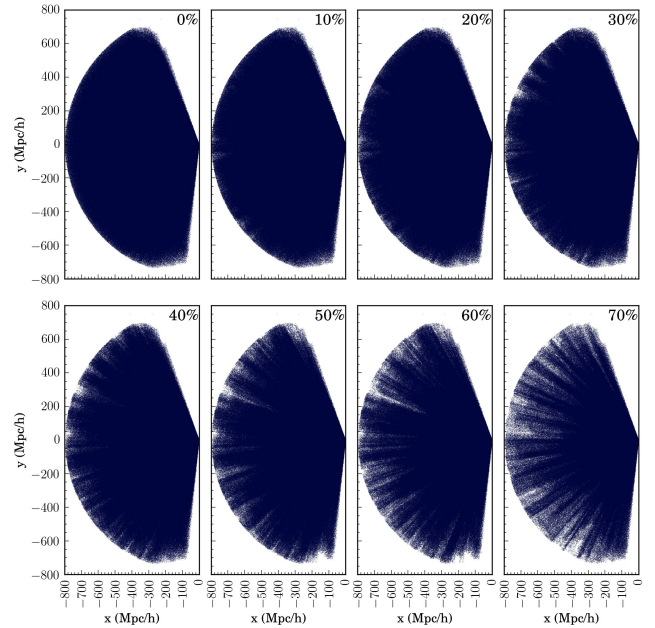
where R_d corresponds to the maximum radial distance from the sample, and $\delta\theta = \min(\alpha_2 - \alpha_1, \delta_2 - \delta_1)$, where α_2, α_1 are the limits in right ascension, and δ_2, δ_1 are the limits in declination. The distortions in redshift space due to peculiar movements may be important for low-redshift galaxy samples. For synthetic catalogues, this effect can be neglected given the nature of the radial selection function, which was built from a statistically homogeneous and isotropic random field in the BOSS footprint. In Fig. 3, a projection in comoving Cartesian coordinates is shown for different samples used in our analysis.

For each synthetic sample, we calculated the generalised correlation integral C_q , the fractal dimension D_q , and the lacunarity spectrum Φ_q . The statistical uncertainty of the C_q values was determined using the equation (4) for the error propagation.

$$\Delta C_q(r) = \frac{1}{MN} \sum_{i=1}^M (q-1) [n_i(r)]^{q-2} \Delta n_i(r). \quad (10)$$

The $\Delta n_i(r)$ uncertainty is determined numerically by counting the number of points that could have been included in $n_i(r)$ but were omitted because of the uncertainty in the position, that is, those points that statistically are outside the spherical shell, despite being very close to its borders ([Chacón-Cardona & Casas-Miranda 2012](#)).

The distance step for the C_q calculation is $\Delta r = r_i - r_{i-1} = 1.0 \text{ Mpc/h}$; this value corresponds to the average distance between two galaxies in clusters ([Narlikar 2002; Sharan 2009; Martínez & Saar 2010](#)). The value of C_q was determined for each sample using 13 structure parameters in the range $-6 \leq q \leq 6$ in order to compare the results with those of ([Chacón 2014; Scrimgeour et al. 2012](#)). Fig. 4 shows the behaviour of C_q with respect to the comoving distance; $q \geq 0$ indicates high-density regions within the distribution, whereas $q < 0$ indicates low-density regions in the same distribution. To avoid divergence in equation (6) when $q = 1$, C_q is calculated using the numerical limit, in which case it converges to the average of the left- and right-hand limits of

**Figure 3.** Redshift-limited samples in comoving Cartesian coordinates for different hole percentage estimates on the BOSS footprint.

q .

We determine the fractal dimension D_q using equation (7). The sliding window technique presented by ([Martínez & Saar 2003; Rodrigues et al. 2004](#)) was applied to the logarithm of the correlation integral as a function of the logarithm of the comoving distance r , because in this case we have a linear relationship between these variables ([Pietronero & Labini 2004; Chacón-Cardona & Casas-Miranda 2012](#)), i.e., $\log(C_q) \propto \log(r)$, where the constant of proportionality is related to the fractal dimension; see equation (5). Thus, the slope $\frac{1}{q-1} D_q$ and the intersection $\frac{1}{q-1} \log(F_q)$ approached a segment of a curve. Fig. 5 shows the fractal dimensions of low- and high-density regions with the structure parameter in the range $-6 < q < 6$. From these plots, it is possible to obtain the behaviour of D_q in terms of the structure parameter, as shown in Fig. 6.

Our analysis is complemented by the lacunarity spectrum, which explain the hole distribution and the effect of border holes, and indicates how the set of points fills the space of the masks. Fig. 7 presents the results for these quantities. In addition, we determined the homogeneity scale for the distribution of points with a margin of error of 1% (see Fig. 8). The dependence of the homogeneity scale r_H on the density of holes in each sample is shown in Table 3.

5 DISCUSSION

According to our results, the spectrum of the fractal dimension provides sufficient information to show a transition or rupture of the homogeneity scale at a specific scale. This justifies the use of statistically homogeneous samples following the observed radial selection function and geometric features of the galaxy masks. Our calculations take into account the

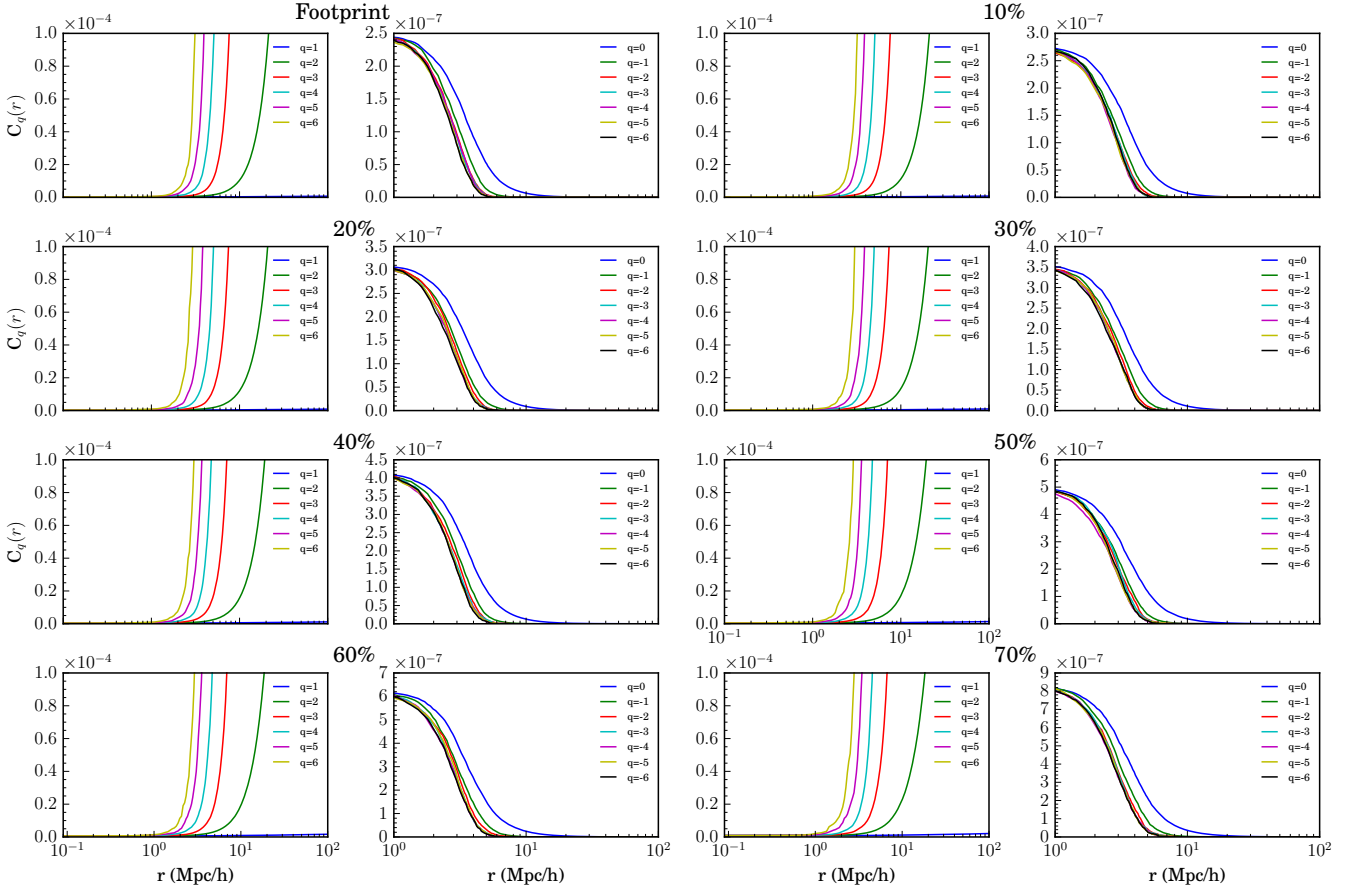


Figure 4. Generalised correlation integral $C_q(r)$ for all the values of q studied in this paper for the footprint samples: left, high densities ($q \geq 1$); right, low densities ($q < 1$).

Table 3. Homogeneity scale r_H for $-6 \leq q \leq 6$ in each sample.

Hole percentage	$q [(r_H \pm 0.03) \text{ Mpc/h}]$					
	-6	-5	-4	-3	-2	-1
0	23.64	20.77	19.11	17.23	17.06	18.95
10	154.58	155.28	150.11	148.65	151.16	144.86
20	73.10	78.85	72.39	71.75	56.49	69.28
30	-	-	64.38	47.69	37.47	67.55
40	87.33	85.49	84.85	79.27	76.69	53.33
50	-	52.19	51.88	54.36	62.16	67.04
60	-	-	-	-	159.25	149.48
70	-	-	-	-	-	-

Hole percentage	$q [(r_H \pm 0.03) \text{ Mpc/h}]$					
	1	2	3	4	5	6
0	12.40	13.70	18.60	16.67	26.15	32.34
10	63.72	68.08	65.49	71.89	72.54	79.36
20	70.61	70.05	68.44	68.07	69.46	67.04
30	73.53	79.85	81.81	81.28	83.25	86.46
40	-	-	-	-	-	-
50	75.21	77.42	81.14	83.13	83.74	88.48
60	93.67	98.87	112.15	116.65	125.60	-
70	88.14	91.98	90.02	93.58	92.34	95.31

maximum radius of the comoving circumscribing sphere of the edges for each sample; therefore, the fractal study can be performed by small and successive iterations within this sphere (Joyce et al. 2005).

The method used to determine the fractal quantities was validated using a main sample without holes, which represents an ideal sample based on the BOSS footprint with 3,273,548 points distributed uniformly across the mask limited to $0.002 < z < 0.2$. In this sample, fractal behaviour does not occur; for all values of the parameter structure, the fractal dimension D_q converges rapidly to the dimension of physical space, $D = 3$, and the transition to homogeneity occurs at $r_H \sim 19.436 \text{ Mpc/h}$.

The fractal dimension $D_q(r)$ was calculated from the generalised correlation integral $C_q(r)$ using the sliding window technique for radial depths that reach as large as 150 Mpc/h . Fig. 4 is consistent with the results of (Chacón 2014; Sarkar et al. 2009); this set of integrals shows the change in the average number of neighbours around the centers, so they increase for $q > 1$ and decrease for $q < 1$. For some q values, these functions indicate the probability that a pair of points are at a distance less than or equal to r . All the samples studied exhibit the same behaviour for C_q independent of the hole density; thus, the fractal dimension will be convergent and will be similar for different values of q . For distances $r > 10 \text{ Mpc/h}$, $C_q \rightarrow 0$, which means that at small

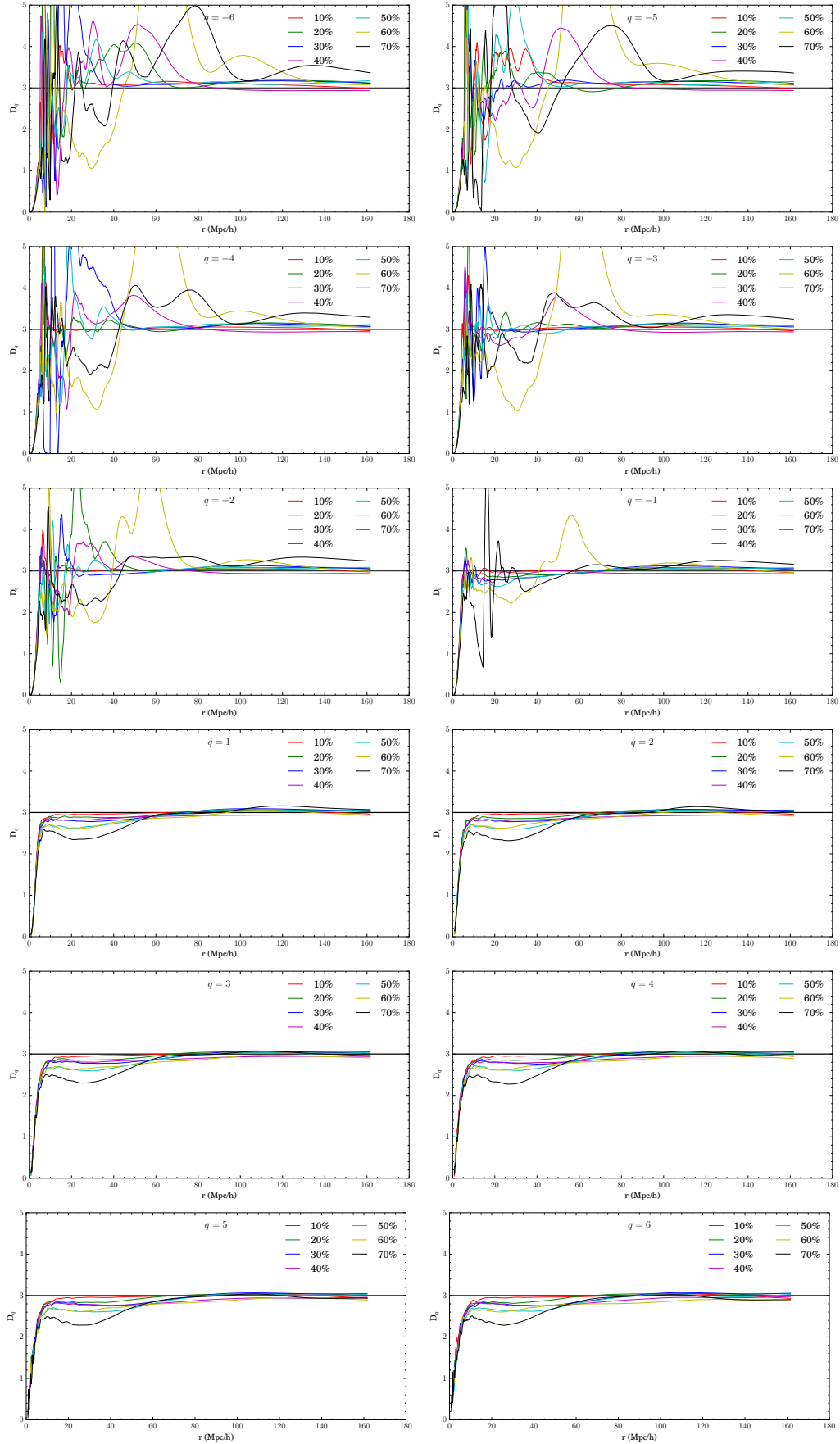


Figure 5. Multifractal spectrum $D_q(r)$ as a function of the comoving radial distance r . For low-density environments, $-6 \leq q \leq 0$; for high-density environments, $q \geq 0$. Error bars have an uncertainty of $\pm 1\sigma$ and coincide with the width line of the curves.

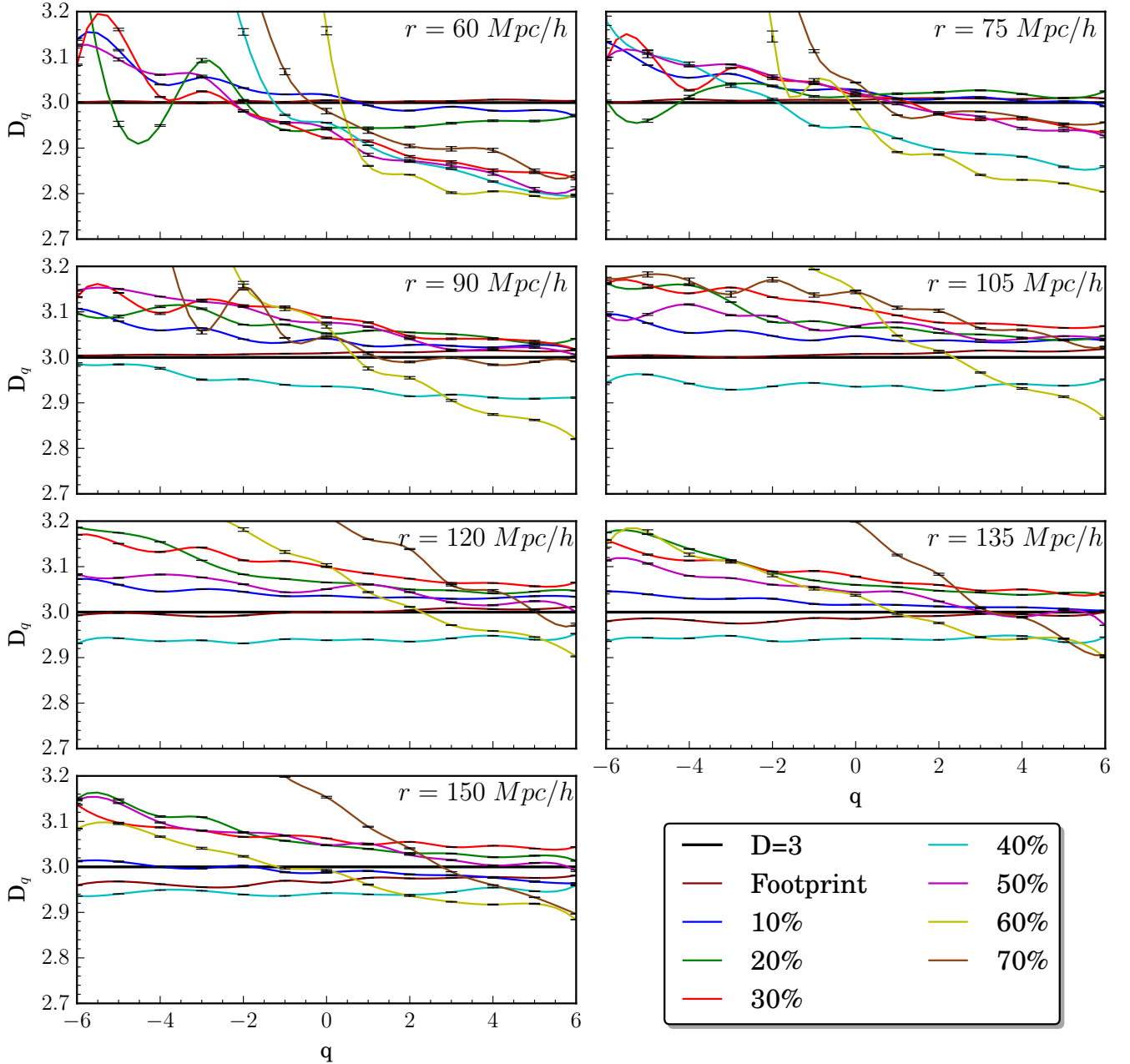


Figure 6. Multifractal spectrum $D_q(r)$ for each sample with structure parameter $q \in [-6, 6]$. The distance r measured from the centers is shown in each subplot. The fractal dimension is bounded by $2.9 \leq D_q \leq 3.2$ and converges to the physical dimension as q increases. The sample with 40% holes exhibits weak fractal behaviour with dimension $2.9 < D_q < 3.0$ at $r > 90 \text{ Mpc}/h$ around the centers in this sample. Solid line shows the Bezier interpolation of the data.

scales, defined structures appear that fill the space in a granular manner because the number of neighbours is less than $q \geq 0$.

The effect of borders is evident in the lacunarity spectrum; this quantity reveals how the space is filled in terms of the hole density and sample morphology, which is given only by the irregular borders of the holes (Gefen et al. 1983; Allain & Cloître 1991). The relative maxima and minima are observed in Fig. 7; these points indicate the distance at which there is a strong dependence on the density of holes. In addition, the distance between peaks is related to the size of

the holes. Φ_q is proportional to the density of holes; for all values of the structure parameter, the lacunarity spectrum presents oscillatory behaviour consistent with the result of (Chacón 2014). The minimum values of Φ_q are between 40 and $120 \text{ Mpc}/h$; that is, there are fewer holes at these distances, and the matter is distributed more evenly than on other scales.

The behaviour of the spectrum of the fractal dimension D_q depends on the radial distance. In low-density regions, D_q fluctuates around the dimension of physical space. For $D_{q<0}$, the dimension tends to $D = 3$ as q increases, indicating that

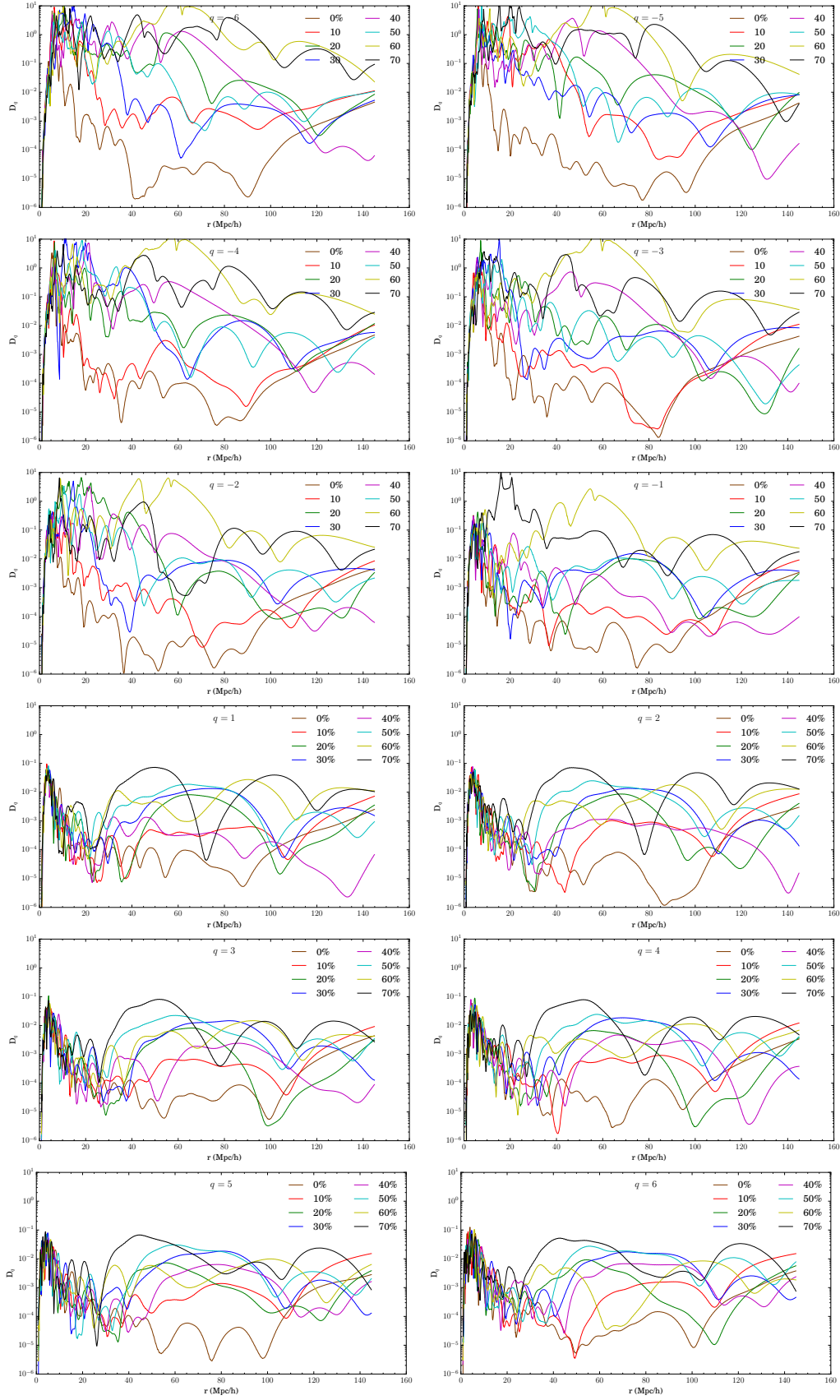


Figure 7. Lacunarity spectrum $\Phi_q(r)$ in logarithmic scale as a function of comoving radial distance r , in low-density environments ($-6 \leq q \leq 0$) and high-density environments ($0 \leq q \leq 6$) for each redshift-limited sample.

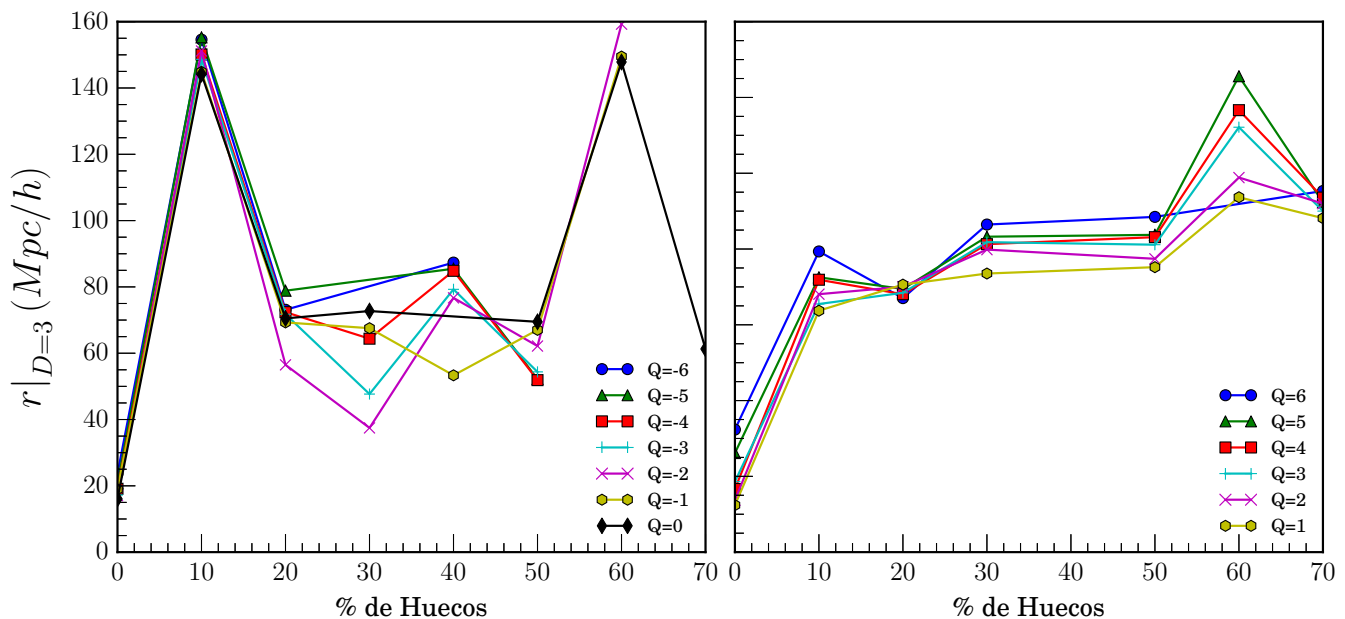


Figure 8. Comparison of homogeneity scale r_H as a function of the hole density: left, low-density environments; right, high-density environments.

the distribution is grouped into small spatial regions with self-similarity properties. When $q < 0$, the amplitude of the fluctuations is large; this causes the growth rate of the fractal dimension to converge to the homogeneity scale r_H . In high-density regions, there is a strong tendency to homogeneity because the values of the fractal dimension are very close of the physical space dimension: $2.97 \leq D|_{r_H} \leq 3.03$. $D_{q \geq 0}$ increases for large r values to reach homogeneity; this means that on average, the space is completely filled at greater depths than r_H . Finally, for $r > 80 \text{ Mpc/h}$, the fractal behaviour disappears.

Our results show a relationship between the hole density from the masks and r_H . In particular, observational holes cause shifts in the homogeneity scale. When the hole density is lower but comparable to the density of objects in the sample, at least for a hole percentage of 40%, a fractal behaviour occurs with $2.70 < D_{q \geq 0} < 3.20$ before the homogeneity scale is reached. For $q > 0$ and hole percentages near 40%, r_H shifts to scales on the order of 120 Mpc/h . For percentages between 10% and 30%, $r_H \sim 70 - 90 \text{ Mpc/h}$, and for percentages below 10%, $r_H \sim 20 \text{ Mpc/h}$, which is very close to the r_H value for the BOSS footprint without holes. This value depends exclusively on the density of points in the samples and the radial selection function. For an infinitely dense sample, the homogeneity scale is $\sim 1 \text{ Mpc/h}$, which is the accuracy of the distance at which sampling is performed.

6 CONCLUSIONS

In this paper, we analysed the multifractal behaviour of synthetic samples built using the BOSS footprint and including different distributions of observational holes and borders. We conclude that the lacunarity increases if the borders describing complex curves that depend on the hole distribution

in the masks. The relative minimum of the lacunarity spectrum in high-density environments ($q \geq 1$) makes it possible to detect the spatial regions where there is a strong dependence on the density of holes. The minimum values of $\Phi_{q > 0}$ are between 40 and 120 Mpc/h ; in these regions, the matter is distributed more evenly than on other scales. In particular, for all the samples with holes and for each $q > 0$, the values of the scale at the homogeneity transition are around 80 Mpc/h .

Synthetic samples without observational holes exhibit a transition to homogeneity around $r_H \sim 19.5 \text{ Mpc/h}$ for all structure parameters. The analysis of samples with holes revealed that for low scales with $q < 0$, there is a fractal behaviour, where fluctuations around the dimension of the physical space are smoothed to achieve homogeneity at $D_q = 3.00 \pm 0.03$ for $r > 130 \text{ Mpc/h}$. This behaviour depends exclusively on the holes and borders in the sample. For $q > 0$, the fractal dimension D_q converges rapidly to r_H , with low lacunarity values and a tendency to fill the physical space on larger scales than 80 Mpc/h .

For the samples studied, we found that r_H depends on the density of observational holes. The holes cause two effects in the calculation of the fractal dimension: shifts in the homogeneity scale proportional to the number density of holes and weakly fractal behaviour with $2.9 < D_q < 3.0$ for a sample with about 40% holes (see Fig. 6), where the hole distribution begins to be more important than the point distribution, so the homogeneity scale disappears slowly for all q values. In this case, the area covered by holes is smaller but is comparable to the area covered by objects. For high-density regions, with $q > 0$ and a percentage of holes near 40%, r_H is shifted to scales on the order of 120 Mpc/h . For areas with 10% and 30% holes, the homogeneity scales are around 70 and 90 Mpc/h , respectively. Similarly, for lower percentages of holes, the homogeneity scale tends to $\sim 19.5 \text{ Mpc/h}$, which corresponds to the r_H value for the

BOSS footprint. Finally, the observational hole percentage in the masks must be taken into account in analyses of the homogeneity scale measured from galaxy samples.

ACKNOWLEDGEMENTS

JEGF acknowledges financial support from the Facultad de Ciencias at Universidad Nacional de Colombia - Sede Bogotá to perform master studies. The authors thank Dr. César Alexander Chacón for providing the code for calculating fractal quantities and Dr. Florian Beutler for suggestions about BOSS footprint.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

REFERENCES

- Allain C., Cloitre M., 1991, *Physical review A*, 44, 3552
- Bagla J., Yadav J., Seshadri T., 2008, *Monthly Notices of the Royal Astronomical Society*, 390, 829
- Baryshev Y. V., Labini F. S., Montuori M., Pietronero L., Teerikorp P., 1998, *Fractals*, 6, 231
- Blanton M. R., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2001, arXiv preprint astro-ph/0105535
- Blanton M. R., Lin H., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2003, *The Astronomical Journal*, 125, 2276
- Blumenfeld R., Mandelbrot B. B., 1997, *Physical Review E*, 56, 112
- Celerier M.-N., Thieberger R., 2005, arXiv preprint astro-ph/0504442
- Chacón C. A., 2014, PhD thesis, Universidad Nacional de Colombia - Sede Bogotá, Departamento de Física, Colombia
- Chacón-Cardona C. A., Casas-Miranda R. A., 2012, *Monthly Notices of the Royal Astronomical Society*, 427, 2613
- Chacón-Cardona C., Casas-Miranda R., Muñoz-Cuartas J., 2016, *Chaos, Solitons & Fractals*, 82, 22
- Conde-Saavedra G., Iribarrem A., Ribeiro M. B., 2015, *Physica A: Statistical Mechanics and its Applications*, 417, 332
- Dawson K. S., et al., 2013, *The Astronomical Journal*, 145, 10
- Durrer R., Labini F. S., 1998, arXiv preprint astro-ph/9804171
- Falconer K., 2004, *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons
- Gabrielli A., Labini F. S., Joyce M., Pietronero L., 2006, *Statistical physics for cosmic structures*. Springer
- Gefen Y., Meir Y., Mandelbrot B. B., Aharony A., 1983, *Physical Review Letters*, 50, 145
- Grujić P., Panković V., 2009, arXiv preprint arXiv:0907.2127
- Hamilton A., Tegmark M., 2004, *Monthly Notices of the Royal Astronomical Society*, 349, 115
- Hogg D. W., Eisenstein D. J., Blanton M. R., Bahcall N. A., Brinkmann J., Gunn J. E., Schneider D. P., 2005, *The Astrophysical Journal*, 624, 54
- Joyce M., Montuori M., Labini F. S., 1999, *The Astrophysical Journal Letters*, 514, L5
- Joyce M., Labini F. S., Gabrielli A., Montuori M., Pietronero L., 2005, *Astronomy & Astrophysics*, 443, 11
- Kazin E. A., et al., 2010, *The Astrophysical Journal*, 710, 1444
- Labini F. S., Vasilyev N. L., Pietronero L., Baryshev Y. V., 2009, *EPL (Europhysics Letters)*, 86, 49001
- Laporte C. F., 2014, PhD thesis, lmu
- Mandelbrot B., 1982, *New\ fork: Freeman*
- Martínez V. J., 1991, in , *Applying Fractals in Astronomy*. Springer, pp 135–159
- Martínez V. J., Jones B. J., 1990, *Monthly Notices of the Royal Astronomical Society*, 242, 517
- Martínez V. J., Saar E., 2003, in , *Statistical Challenges in Astronomy*. Springer, pp 143–160
- Martínez V. J., Saar E., 2010, *Statistics of the galaxy distribution*. CRC Press
- Munoz-Cuartas J. C., Mueller V., 2012, arXiv preprint arXiv:1203.4846
- Murante G., Provenzale A., Spiegel E., Thieberger R., 1997, *Monthly Notices of the Royal Astronomical Society*, 291, 585
- Narlikar J. V., 2002, *An introduction to cosmology*. Cambridge University Press
- Peebles P. J. E., 1980, *The large-scale structure of the universe*. Princeton university press
- Peebles P., 1989, *Physica D: Nonlinear Phenomena*, 38, 273
- Peebles P. J. E., 1993, *Principles of physical cosmology*. Princeton University Press
- Pietronero L., Labini F. S., 2004, in , *Advances in Solid State Physics*. Springer, pp 375–388
- Pino E., Hetem A., Horvath J. E., de Souza C. A., Villela T., De Araujo J., 1995, arXiv preprint astro-ph/9502048
- Rodrigues E. P., Barbosa M. S., Costa L. d. F., 2004, arXiv preprint cond-mat/0407079
- Ross A. J., et al., 2012, *Monthly Notices of the Royal Astronomical Society*, 424, 564
- Sarkar P., Yadav J., Pandey B., Bharadwaj S., 2009, *Monthly Notices of the Royal Astronomical Society: Letters*, 399, L128
- Saslaw W. C., 2000, *The distribution of the galaxies: gravitational clustering in cosmology*. Cambridge University Press
- Scrimgeour M. I., et al., 2012, *Monthly Notices of the Royal Astronomical Society*, 425, 116
- Seshadri T., 1999, *Pramana*, 53, 989
- Seshadri T., 2005, *Bulletin of the Astronomical Society of India*, 33, 1
- Sharan P., 2009, *Spacetime, geometry and gravitation*. Vol. 56, Springer Science & Business Media
- Smee S., et al., 2012, arXiv preprint arXiv:1208.2233
- Springel V., et al., 2005, *nature*, 435, 629
- Wu K. K., Lahav O., Rees M. J., 1999, *Nature*, 397, 225
- Xia X.-Y., Deng Z.-G., Zou Z.-L., 1992, *Science in China A: Mathematics*, 35, 326
- Yadav J., Bharadwaj S., Pandey B., Seshadri T., 2005, *Monthly Notices of the Royal Astronomical Society*, 364, 601
- Yadav J. K., Bagla J., Khandai N., 2010, *Monthly Notices of the Royal Astronomical Society*, 405, 2009
- Zheng W., Zu-gan D., Yong-zhen L., yang Xia X., 1988, *Chinese astronomy and astrophysics*, 12, 269

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.